Understanding the role of firms in the gender wage gap over time, over the life cycle and across worker types[*]

Marco G. Palladino[†] Alexandra Roulet[‡] Mark Stabile[§]

March 15, 2022

Abstract

We revisit evidence on the contribution of firms to the gender wage gap using a cluster-based approach to investigate time series and life-cycle patterns as well as match effects by gender. This also relaxes usual sample restrictions, resulting in larger estimates of the contribution of firms, driven by a higher within-firm component. Further, despite a decline in the unconditional gender wage gap between 1995 and 2015, the gap in firm pay premiums and its decomposition remained constant. It increases with age, exclusively driven by the between-firm component. Finally we find limited evidence of complementarities for both men and women.

Keywords: gender wage gap, firm pay premium, sorting, bargaining, discrimination

JEL Codes: J16, J31, J71

[*] We would like to thank Pierre Cahuc, Nicolò Dalvit, Kerstin Holzheu, Barbara Petrongolo, Roland Rathelot, Jean-Marc Robin and in particular Stéphane Bonhomme, Francois Gerard, Thibaut Lamadon, Christian Moser and Raffaele Saggio for their helpful comments. We thank seminar participants at Queen Mary University, Stockholm University, Princeton, Banque de France and Zurich UZH as well as participants at the 4th CEPR-Bank of Italy Labor Workshop, the Junior Micro Macro Labor Conference at Columbia University, INSEAD Brown Bag and Sciences Po PhD Lunch. The technology to access the data remotely is supported by a public grant overseen by the French National Research Agency (ANR) as part of the 'Investissements d'avenir' programme (ANR-10-EQPX-17 – Centre d'accès sécurisé aux données – CASD).

[†] Department of Economics, Sciences Po. marco.palladino@sciencespo.fr

[‡] INSEAD and CEPR. alexandra.roulet@insead.edu

[§] INSEAD and CEPR. mark.stabile@insead.edu

1 Introduction

Understanding what drives differences in wages between women and men has been the subject of considerable economic research for decades (Blau and Kahn, 2017; Olivetti and Petrongolo, 2016). Recently, research examining the role that firms play in contributing to the gender wage gap aims to separate two important drivers: the extent to which women sort into lower paying firms and the extent to which women extract less surplus than men within a given firm. In particular the seminal work of Card et al. (2016) has adapted the Abowd et al. (1999) (AKM from now on) framework to assess the impact of firm-specific pay policies on the gender wage gap. This methodology estimates wage regressions with both worker and firm fixed effects separately by gender in order to compute gender-specific firm effects which allows the researcher to compute a gender gap in firm pay premiums and to decompose it into a sorting or between firm channel (what women would be paid if they worked at the firms where men work) and a bargaining or within firm one (what women would be paid at these firms if they were men).

The empirical application in Card et al. (2016) (CCK from now on) uses Portuguese data, but several papers have since then leveraged this methodology to study other contexts (Bruns, 2019; Casarico and Lattanzio, 2019; Coudin et al., 2018; Gallen et al., 2019; Morchio and Moser, 2020; Sorkin, 2017). Although magnitudes vary across countries, results consistently show that sorting between firms explains most of the firm component of the gender wage gap (see Table 1 for a summary of selected estimates). In other words, gender differences in pay premiums are largely driven by men and women working in different firms rather than similar men and women being paid differently at the same firm.

While the gender gap in firm pay premiums is therefore well documented, there is less evidence on how this gap, and its decomposition into between and within firm differences, have evolved over time and change over the life cycle. There is also little evidence on whether men and women benefit differently from match effects i.e. from complementarities between their type and their firm's type.

To shed light on these research questions, we leverage new insights from Bonhomme et al.

(2019) (BLM from now on). Instead of estimating firm effects for each firm, they propose to first group or cluster firms that are sufficiently similar and then estimate cluster effects. Reducing the dimensionality of the estimation has several nice properties, some of which are particularly useful in our context. First, this clustering approach does not require the sample restriction imposed by the AKM methodology. In a specification with both firm and worker fixed effects, firm effects are only identified for firms that are connected by worker mobility and *gender specific* firm effects are only correctly identified within the "dual connected set", the intersection of the male and female largest connected sets. In contrast the cluster approach allows us to include all firms and group them into clusters. Using clusters also makes it easier to estimate very short panels. The AKM specification is not as suited for short panels because of the sample restriction just outlined: the shorter the panel, the fewer the movers observed and the more restrictive the estimation sample. In contrast the cluster approach is well suited for 2 years panels. Moreover, it also makes it easier to estimate effects on precise age brackets: again, with an AKM specification, restricting to a given cohort or age bracket results in too few movers, whereas this is not an issue with clusters. Finally, reducing the dimensionality of the fixed effects allows to account for complementarities between workers and firms in a more tractable way [1].

We therefore exploit the clustering approach proposed by BLM to estimate the evolution of the contribution of firms to the gender wage gap over time using repeated 2 years panels and over the life cycle using age-specific estimates, in both cases in the absence of dual connected set restriction. We also explore, using their static interaction model, whether match effects/complementarities exist in the French context and, in particular, whether they differ for men and women.

We first show that lifting the sample restriction of the dual connected set leads to higher estimates of the gender gap in firm pay premiums driven by a higher within component, while the AKM and BLM methods produce a very similar decomposition when restricting to the same sample. Further, we show that the contribution of firms to the gender wage

[1] It also allows for the modelling of some dynamic aspects of the wage setting, which we abstract from in this paper. See Di Addario et al. (2021) for a dynamic extension of the AKM two-way fixed effects model and Costa Dias and Holzheu (2021) for a more general dynamic specification for wages in the gender context.

gap has increased over the last 2 decades. In absolute terms, from 1995 to 2015, the difference between the average firm effect for males and the average firm effect for females has remained quite stable around 3 to 4 log points, with 2 log points due to the between firm gap and 1 to 2 log points to the within firm one. But in 1995 it represented just above 15% of the unconditional gender gap in hourly wage while in 2015 it was 30%. Moreover we show that the gender gap in firm pay premiums increases over the life-cycle, exclusively driven by an increase of the sorting of women into lower paying firms. Finally we find little evidence of complementarities for both men and women, which suggests that in our context an additive specification without interaction terms between worker and firm types sufficiently captures the main effects we focus on.

The rest of the paper proceeds as follows: Section 2 describes the data. Section 3 presents the estimation framework, and in particular how we cluster firms together. Section 4 shows the results, and section 5 concludes.

2 Data

Our data come from matched employer-employee registers in France (DADS data). Our main dataset is the DADS Panel which allows us to follow a subset of workers from 1995 to 2015. It provides information on the firm in which they are employed as well as their earnings and other administrative data on their employment. The sampling rate is 1/12th from 2002 onwards and 1/24th before. We discard all public sector, temp agency and self employed workers and restrict the sample to workers who are between 25 and 60 years old and who have worked at least half the year (summing across all their jobs)[2]. If workers have multiple jobs in a given year, we use the information from the highest paying job.

We have information on hours worked such that our main wage measure is the hourly wage. In France the unconditional gender gap in hourly wage went from 19% in 1995 to 14% in 2015. Table 2 describes our sample for the period 2009-2015. On average, men and women are 40 years old. Women earn on average 26,000 euros gross per year while for

[2]These restrictions are made to drop workers too weakly attached to the labor market or who have non-standard labor contracts.

men it is 32,000 euros. 27% of women are part-time and 7.5% of men.

We also make use of the DADS Postes dataset, which gives us information on the universe of jobs for the universe of firms every year. We use these data to cluster firms together as described in section 3.1. Since firms are clustered based on their empirical cdf of wages, it is important to observe the full distribution of wages within each firm to accurately partition firms into clusters. However, once we have the mapping between each firm identifier and its cluster, because the DADS Postes dataset is a repeated cross-section from the point of view of workers (their identifier changes every year), we then use the DADS Panel dataset above-mentioned to run regressions with worker fixed effects (the firm identifier is the same in DADS Postes and DADS Panel datasets, allowing us to merge the firms' assignment to clusters obtained from DADS Postes to the DADS Panel data). Following BLM, we drop a few sectors (fishery and agriculture, education, health, and social work).

3 Estimation framework

To identify the contribution of firms to the gender wage gap, we begin with an estimating equation similar to that used in CCK except that we estimate gender-specific cluster effects instead of gender-specific firm effects in order to eliminate the sample restrictions required in CCK and to allow us to document time trends and changes over the life cycle by estimating effects on respectively overlapping short panels and small age brackets. We then build on this estimating equation by adding interaction terms between worker types and clusters, following the BLM static interaction model, to explore whether there are complementarities and if they differ by gender.

3.1 Clustering firms

The BLM framework partitions firms into clusters based on the similarity of their earnings distributions. Cluster fixed effects then replace individual firm fixed effects.

Specifically, the clusters are obtained by solving the following weighted k-means problem:

$$\min_{k(1),..,k(J),H_1,..,H_K} \sum_{j=1}^{J} n_j \int \left(\widehat{F}_j(y) - H_{k_j}(y)\right)^2 d\mu(y) \tag{1}$$

where n_j is the average number of workers of firm j, μ corresponds to the discrete support of the cdf (in our case deciles), $k(1), .., k(J)$ denotes a partition of firms into K known classes and $H_1, .., H_K$ are generic cdfs. The minimization is with respect to all possible partitions and class-specific cdfs. In other words, for each possible partition of firms into K groups, we can compute for each firm the distance between its log earnings cdf $\widehat{F}$ and the centroid cdf H of the group to which the firm is assigned. Then for each partition we can sum across all firms these distances (squared). The algorithm picks the partition which minimises this sum.

3.2 What do clusters capture?

Grouping firms into clusters has several advantages. First, though not central to our focus on the gender wage gap, it is a way to estimate firm effects without being subject to limited mobility bias. When the wage is regressed on a full set of firm and worker dummies, the firm effects are only identified from the workers moving across firms and if there are only a few movers then this creates bias. Clusters get around this problem because there are a sufficient number of workers who move between clusters to identify the cluster fixed effects. It has been shown that this limited mobility bias leads to an overestimate of the share of the variance of earnings explained by firms and an underestimate of the share explained by the covariance between worker and firm effects (Andrews et al., 2012)[3]. However, in our context where we are exploring differences in average firm effects by gender, the AKM estimator should be unbiased on average, even if noisy.

More relevant for our purposes, clusters allow us to lift the sample restriction of the dual connected set. As mentioned above, comparing male and female firm effects in the absence

[3] See variance decomposition of wages in our data using firms v. clusters in Appendix Table A1. We also replicate an exercise proposed by Andrews et al. (2012) and present the results in Figure A1. We find that the correlation between individual and firm/cluster fixed effects is increasing in the average number of movers when we use AKM and stable when using clusters.

of clusters requires restricting the sample to firms with both male and female movers in order to identify a firm effect for both genders, more precisely to the intersection of the male and female largest connected sets. Because of this, small firms will be underrepresented and single gender firms will be excluded. A variant of firm and worker fixed effect models consists of treating all firms with less than 10 employees as if they were a single firm (one single fixed effect for all these firms). Although this variant avoids dropping small firms, it still requires having movers of both genders for firms of more than 10 employees and it imposes a single estimate for all small firms, which might be inaccurate. Instead, when using clusters all firms are included and treated in a consistent manner. As we don't have to worry about this sample restriction, which is more acute the shorter the length of the panel, we can estimate our models on repeated, overlapping very short panels in order to document time trends. Similarly, we can estimate not only gender-specific but gender-*and age-specific* estimates of pay premiums in order to document the evolution over the life-cycle. A third nice feature of clusters for our purpose is that it allows to estimate match effects (complementarities) in a tractable way, as we discuss below.

Clustering also raises some concerns. For example, are clusters more informative than just a set of firm characteristics such as sectors or firm size? Figures A2 and A3 show that there is substantial sectoral variation within clusters as well as variation in firm size, suggesting that cluster fixed effects capture something different than just firm size or industry/sector fixed effects.

3.3 Estimating the contribution of firms to the gender wage gap with clusters

As mentioned earlier, to identify how much firms contribute to the gender wage gap, we use an estimating equation similar to CCK except using gender-specific cluster effects instead of gender-specific firm effects. More precisely, we define the wage of worker i in year t, whose gender is $G(i) \in \{F, M\}$ and who is currently employed in firm $J(i,t)$, as the sum of a person fixed effect, a cluster fixed effect, other controls and a residual error term:

$$Log(w_{it}) = \alpha_i + \sum_{k=1}^{K} \Psi_k^{G(i)} \mathbb{1}(J(i,t) = k) + \beta^{G(i)} X'_{it} + r_{it} \quad (2)$$

α_i is the worker fixed effect. X_{it} are time-varying controls. In our baseline specification we control for a polynomial in age, experience and its square, tenure at the firm (since 1995) and year dummies. For robustness we also show results when we control as well for occupations, which has little effect on the estimates[4]. $\mathbb{1}(J(i,t) = k)$ are dummies indicating the cluster to which firm $J(i,t)$ has been assigned. $\Psi_k^{G(i)}$ is the gender-specific effect of being in such a cluster. Indeed we run these regressions separately for men and women so as to collect *gender-specific* cluster effects, which capture time-invariant factors that impact the average surplus for all employees of a given gender in a given cluster: $\Psi_k^{G(i)} \equiv \gamma^{G(i)}\overline{S}_k$. These are our main objects of interest.

Finally, r_{it} is a composite error term. In our base specification it includes the individual's idiosyncratic error as well as the time-varying and match specific effects of being in a given cluster: $r_{it} = \gamma^{G(i)}(\phi_{k,t} + m_{i,k}) + \epsilon_{it}$. However, by reducing the dimensionality of the firm fixed effects with the cluster approach, we will be able to account for complementarities in subsequent analyses below.

3.4 Normalization

Whether using firm dummies or cluster dummies, the firm/cluster effects are only identified relative to a constant. We thus cannot compare male and female effects without first normalizing them consistently. CCK's original paper and subsequent literature using this method have shown that firm effects are typically similar for low value added firms and then start to increase with value added per worker in a linear way. We illustrate this graphically in Figure 1. This kinked relationship suggests that there is no rent to share below a certain threshold of value added per worker and that pay premiums matter only after that threshold. Thus prior work has computed, for each gender, the employment-weighted average firm effect of low value added firms below this threshold and de-meaned all firm effects by this constant[5]. We proceed similarly. We will highlight later that this only mat-

[4]We do not have access to education but education is mostly fixed over time, hence captured in the worker fixed effect.

[5]This requires restricting the sample to firms with Value Added information, which drops around 15% of firms from the universe. We make this restriction at the very beginning of our analyses (before defining the

ters for the estimates of the within component of the gender gap in firm pay premiums. An alternative normalization, which we show in some robustness exercises, is to de-mean all firm/cluster effects relative to average effects in the hotel and restaurant sector with the hypothesis that this is a sector with no surplus to share.

3.5 Between firm versus within firm decomposition

Once we have estimated cluster effects for men and women separately and normalized them, we can perform the CCK decomposition of firm/cluster effects into a between component (sorting channel) and a within component (bargaining channel).

$$\begin{aligned} E[\Psi_k^M|Male] - E[\Psi_k^F|Female] = E[\Psi_k^M - \Psi_k^F|Male] + \\ E[\Psi_k^F|Male] - E[\Psi_k^F|Female] \end{aligned} \tag{3}$$

The left-hand side of equation (3) is the difference between the average cluster fixed effect for men and the average cluster fixed effect for women. We can then decompose it into two components.

The first term on the right hand side of equation (3) represents the bargaining, or within-cluster, channel. It is the average difference between the cluster fixed effects for men and women, assuming that women are represented in the different clusters in the same proportion that men are, i.e. holding sorting fixed.

The other two terms represent the sorting, or between-cluster, channel. It is the difference between the average cluster effect for women and what it would be if women were represented in the different clusters in the same proportion that men are. If men and women were equally represented in each cluster, this sorting term would be equal to zero.

Note that we can rewrite the formula above holding the distribution of women across clusters as fixed instead of that of men for the within-cluster channel and using the male

clusters or before defining the dual connected set). The threshold (dashed vertical line in Figure 1) is defined more formally as in Card et al. (2016): for each possible threshold τ in the distribution of log of value added per worker (VA_j where j indexes firms), we compute the mean squared error of a system of two equations (one for each gender $G \in (M, F)$), $\psi_j^G = \pi_0^G + \pi^G VA_j + \epsilon_j^G$, when dropping all observations to the left of τ. We keep the τ that minimises the mean squared error of this system of two equations.

premium for the between-cluster channel.

$$E[\Psi_k^M|Male] - E[\Psi_k^F|Female] = E[\Psi_k^M - \Psi_k^F|Female] + \\ E[\Psi_k^M|Male] - E[\Psi_k^M|Female] \tag{4}$$

We prefer decomposition (3) as it reports sorting as what women would be paid if they worked at firms where men worked and the within component as what they would be paid at these firms if they were men (instead, under the alternative decomposition, sorting gives what men would be paid at firms where women work and the within component what they would be paid if they were women, which seems a less natural way to define our objects of interest.). We therefore report mostly results of decomposition (3) in what follows although we always checked robustness to the alternative one.

Since the sorting estimates only use one set of firm effects (either the male or the female ones depending on the decomposition used), normalization does not matter here. The constant will cancel out in the two terms. However, with the within estimates, where the constant does not cancel out, the choice of normalization will affect the estimates.

3.6 Allowing for worker-firm/cluster complementarities

Linear two-way fixed effects regression models as the one in equation (2) do not allow for the possibility that different types of workers benefit differently from firm effects. Yet a large theoretical and structural literature has highlighted that sorting patterns between workers and firms can be partly accounted for by the presence of complementarities in production and wages: workers' idiosyncratic skills may be particularly valuable for certain firms (Becker, 1973; Shimer and Smith, 2000; Eeckhout and Kircher, 2011; Hagedorn et al., 2017). We therefore want to allow for complementarities in our analysis in order to understand whether wage differences between men and women are also explained by *gender-specific* differences in these match effects.

We thus further estimate non linear earnings models which explicitely feature complementarities between firms/clusters and workers. The estimating equation can be written as follows:

$$Y_{i,t} = a(k(i,t)) + b(k(i,t)) \times \alpha_i + \epsilon_{i,t} \tag{5}$$

Where i denotes the individual, t the year, Y is the log hourly wage residualized on controls, and $k(i,t)$ the cluster to which is assigned the firm in which individual i is employed in year t. We closely follow the estimation procedure of Bonhomme et al. (2019) such that the above equation is a finite mixture model where the α_i are random effects and t only takes two values - period 1 and 2 - while the wage is residualized on year dummies and a polynomial in age. With this model, instead of getting a single estimate for each cluster, we estimate two sets of parameters: the $a(k)$ give the cluster effects for workers of type $\alpha_i = 0$, and the $b(k)$ the extra gain for higher type workers. Complementarities refer to differences in $b(k)$ across clusters. Identification of these parameters still come from movers across clusters. We can identify both the $a(k)$ and the $b(k)$ as long as movers back and forth between a given pair of clusters (from cluster k to k' and from cluster k' to k for example) are not of the same type. Indeed as long as $E_{kk'}(\alpha_i) \neq E_{k'k}(\alpha_i)$, we can recover

$$\frac{b(k')}{b(k)} = \frac{E_{kk'}(Y_{i2}) - E_{k'k}(Y_{i1})}{E_{kk'}(Y_{i1}) - E_{k'k}(Y_{i2})}$$

Of course we are interested in gender-specific effects, so we estimate equation (5) separately for each gender, which yields, for each cluster k gender-specific estimates $a^G(k)$ and $b^G(k)$ with $G \in \{M, F\}$. If we want to compare these estimates across gender, we once again need to normalize them consistently. We proceed as already described in section 3.4, taking as reference group low-value added firms. Note, however that this is only required to make quantitative statements about differences in male and female estimates. The assessment of whether there exists, or not, complementarities does not require any specific ex-post normalization. For that, we just check, for each gender, whether the b(k) are similar for each cluster. In a world without complementarities $\forall k, b(k) = 1$, and we are back to a simple additive model.

3.7 Estimating the contribution of complementarities to the gender wage gap

After estimating equation (5) for each gender, we first check, as just mentioned, whether there seems to be complementarities, either for males or females or both. Then we normalize the estimates and try to quantify the contribution of these potential match effects, or complementarities, to the gender wage gap. The difference in the average (residualized) wage between men and women is defined as follows:

$$\begin{aligned}\mathbb{E}[Y_{i,k}|Male] - \mathbb{E}[Y_{i,k}|Female] =& \mathbb{E}[a^M(k)|Male] + \mathbb{E}[b^M(k) \times \alpha_i^M|Male] - \\ & \mathbb{E}[a^F(k)|Female] + \mathbb{E}[b^F(k) \times \alpha_i^F|Female]\end{aligned}$$

By adding and subtracting the term $\mathrm{E}[\alpha_i^M|Male] - \mathbb{E}[\alpha_i^F|Female]$, it can be rewritten as:

$$\mathbb{E}[Y_{i,k}|Male] - \mathbb{E}[Y_{i,k}|Female] = \tag{6}$$

$$\mathbb{E}[a^M(k)|Male] - \mathbb{E}[a^F(k)|Female] + \tag{7}$$

$$\mathbb{E}[\alpha_i^M|Male] - \mathbb{E}[\alpha_i^F|Female] + \tag{8}$$

$$\mathbb{E}[(b^M(k) - 1) \times \alpha_i^M|Male] - \mathbb{E}[(b^F(k) - 1) \times \alpha_i^F|Female] \tag{9}$$

$\mathbb{E}[a^M(k)|Male] - \mathbb{E}[a^F(k)|Female])$ captures the average gender difference in firm pay premiums without accounting for complementarities. $\mathbb{E}[\alpha_i^M|Male] - \mathbb{E}[\alpha_i^F|Female]$ represents the average gender difference in worker unobserved heterogeneity. Finally, $\mathbb{E}[(b^M(k) - 1) \times \alpha_i^M|Male] - \mathbb{E}[(b^F(k) - 1) \times \alpha_i^F|Female]$ can be seen as the contribution of match effects to the gender wage gap. The idea here is to capture how much the estimated set of $b^G(k)$ inflates worker's unobserved heterogeneity, and to see if this differs by gender.

4 Results

4.1 Firm versus cluster effects, holding sample fixed

We first compare results of decomposition (3) when using firm v. cluster dummies, holding sample fixed. We use a 6 year panel from 2009-2015 and restrict attention to observations that are in the dual connected set[6]. On this sample we can estimate the model of equation (2) with either firm or cluster dummies. We normalize the estimates with respect to low value added firms. Table 3 shows that average pay premiums are reasonably similar, between 0.8 and 1.5 log points apart, whether using firm dummies or cluster dummies. This builds confidence in the use of clusters.

Further, Table 4 shows that the gender gap in pay premiums is also similar whether using firm or cluster dummies and, importantly, that the decomposition into between and within components is also quite similar in both cases, alleviating concerns that the within component under a cluster approach would capture some between firm effects. Moreover Table 4 shows that results are very similar when we increase the number of clusters. Based on this evidence, we mostly show results with 10 clusters in what follows, although we confirm that our results are robust to using more clusters. Overall, in France, the gender gap in pay premiums in the dual connected set is around 3 log points with 70 to 75% due to a between-firm effect and 25 to 30% due to a within-firm one.

4.2 Lifting the dual connected set restriction

Table 5 reports results using the full sample compared to limiting the set of firms to those in the dual connected set. As can be seen in Table 2, the number of firms in the full sample is around 300,000 compared to just above 17,000 in the dual connected set. The main difference between the two samples is firm size: in the dual connected set, firms have on average 283 employees, whereas in the full sample this number is 31. Besides that, worker and firm characteristics are fairly similar.

Table 5 shows that in the full sample, compared to the dual connected set: i) the unconditional gender wage gap is lower (this finding is consistent with what is reported in other papers); ii) the pay premium gap is higher; iii) the share of the within component is higher.

[6]We first define the largest connected set on the sample of females and the largest connected set on the sample of males. We then keep firms that are in the intersection of both sets and estimate equation (2) either with firm effects or with cluster dummies, where clusters are defined among firms that are in the dual connected set. We use a six year panel in order to both avoid having a too short panel which would include too few firms in the dual connected set, and to avoid the panel being too long which would potentially invalidate the time invariance hypothesis behind the estimation of firm fixed effects. When we use 2 year panels, we can just cluster firms based on the information of one year. When we use longer panels, such as here, we detrend the wage and cluster based on average values over the period considered.

The unconditional gender gap in hourly wage falls from 17% to 13.5%. The gender gap in firm pay premium increases from 3.1 to 4.2 log points in absolute terms, and more meaningfully as a share of the unconditional wage gap. The between firm component falls as a share of the pay premium gap from 75% to 50% and the within component increases from 25% to 50%.

Figure 2 reports average gender gap in pay premiums and the decomposition into a between and within component by firm size. We use the same estimates as those used in Table 5 for the full sample but we compute separate averages for employees of firms of different size instead of computing averages for the whole population. We define three firm size categories: less than 20 employees, between 20 and 250 and more than 250 employees. We see that the unconditional gender wage gap increases with firm size while the gender gap in firm pay premiums remains roughly stable at 3 to 4 log points. However, the decomposition changes strikingly from being mostly explained by the within component for smaller firms to being mostly explained by the between component for larger firms. Hence lifting the dual connected set restriction, which results in the inclusion of many more small firms, pushes the share of the gender gap in firm pay premiums explained by the within component upwards. A possible explanation for this heterogeneity by firm size may be that larger firms could have more processes in place to ensure equal pay for similar jobs while small firms might have more discretion in pay across employees.

4.3 Time series evidence on the contribution of firms to the gender wage gap

We next present evidence on the evolution of the contribution of firms to the gender pay gap over time. We use the full sample of firms between 1995 through 2015 and estimate equation (2), the gender gap in firm pay premiums and its decomposition into between and within components, separately for each overlapping 2 year panel. Figure 3 shows the results. Despite a decline in the unconditional gender gap in hourly wage between 1995 and 2015 from 19% to 14%, the gender gap in firm pay premiums has remained fairly stable over these two decades : it thus represents an ever increasing share of the gender wage gap. The between component has remained particularly stable around 2 log points while the within component fluctuates between 1 and 2 log points. The panel on the left shows results when using 10 clusters, normalizing with respect to low value added firms and our preferred decomposition from equation (3). The panel on the right shows robustness to adding 2-digit occupations to our set of usual controls (polynomial in age, experience and tenure). Including occupational controls has little effect on our estimates of firm pay premiums suggesting that workers' mobility across firms, which identifies firm effects, does not typically coincide with occupation switches. We also see that the share of the gender wage gap explained by firm pay premiums (red line) is about the same as that explained by occupation (which is more or less the difference between the black and grey lines in the right panel[7]). Appendix Figure A4 shows robustness to normalizing with respect to the hotel and restaurant sector (left panel) and to using 50 clusters (right panel).

4.4 The contribution of firms to the gender wage gap over the life cycle

Another advantage of using clusters is that we are able document the evolution of the contribution of firms to the gender pay gap over the life cycle using age specific estimates. While some of the previous literature using firm effects instead of clusters has examined such heterogeneity, it did so using estimates for the entire population instead of estimates specific to the age bracket being considered. For example, $E[\Psi_k^M | Male\&age = 25-30] - E[\Psi_k^F | Female\&age = 25-30]$ would give the gender gap in firm pay premiums for workers between 25 and 30 years old using the firm effects estimated on the entire sample, not the sample restricted to those 25 through 30. The exercise can be repeated for any age bracket. All the potential age heterogeneity in these gaps comes from whether the distribution of males and females across firms varies by age. In contrast, with clusters, because there are more movers, we can restrict the sample to workers of age 25 to 30 and compute $E[\Psi_k^{M,25-30} | Male\&age = 25-30] - E[\Psi_k^{F,25-30} | Female\&age = 25-30]$. In this equation, the estimates of pay premiums vary not only by gender but also by age. Specifically we can estimate cluster effects following equation (2) on overlapping 5 years age brackets. For each age bracket, we normalize estimates relative to low value added firms. We then average

[7]The difference between the black and grey lines in the right panel also includes the effects of age, experience and tenure but, as can be seen by the difference between the black and grey lines in the left panel, these other controls explain very little of the gender wage gap.

estimates by age and focus on the age range for which we have 5 sets of estimates, i.e. from age 30 to 55. In this exercise, we use our whole sample period 1995-2015 and the same clusters throughout the 20 years. This allows us to document not only the life cycle pattern of the gender gap in firm pay premiums but also its evolution by cohort. Indeed, we can average our estimates not only by age but by age and cohort. Results are consistent if we focus on a more narrow period.

The results are presented in Figure 4. First we observe that the unconditional gender pay gap increases over the first part of the life cycle but remains relatively flat after about age 40 (panel a). The gap has also declined for each subsequent cohort. The gender gap in firm pay premiums (panel b) increases over the life cycle (with a similar plateau at around age 40) and we observe almost no differences across cohorts. Interestingly, there is little evidence that the within component increases over the life cycle or that it has changed substantially across cohort (panel c). It is really the between firm or sorting component that increases over the life cycle, accounting for the rise in the overall firm pay premium gap. This between component has remained similar across subsequent cohorts of workers (panel d). In sum, despite substantial age heterogeneity, there is much less cohort heterogeneity in firm effects consistent with the evidence on changes over time presented in the previous sub-section[8].

4.5 Evidence on complementarities for both men and women

We next investigate whether there is evidence of complementarities between firm and worker types and whether these differ by gender. We first residualize the wage on year dummies and a polynomial in age and then estimate equation (5) with this residualized wage, separately for each gender[9]. We use the same clusters as those used in Table 3

[8] Note that instead of computing estimates by age-bracket and averaging them by age and cohort, we could compute cohort-specific estimates and average them by age (and cohort). The results are presented in Appendix Figure A5 and consistent with the results of Figure 4.

[9] Our analysis uses the code made available online by BLM, except that we define the clusters ourselves in order to cluster using information on the entire wage distribution in the universe of firms from the DADS Postes, and not just the information from the panel of workers that we can follow over time (DADS panel, see section 2 for more details). We define the movers and stayers for each 2 years sub-period between 2009 and 2015 and append the resulting datasets into one mover and one stayer datasets. The variance decomposition associated to this analysis is shown in Table A2.

through 5 for the 2009-2015 period.

The results are presented graphically in Figure 5. Following BLM, even though worker effects are estimated as continuous random effects, we show results graphically for a discrete number of worker types. Specifically the distribution of worker effects is split into 6 equal size bins. The figure reports for each gender $G \in \{M, F\}$, worker type $l \in [1, 6]$ and cluster $k \in [1, 10]$, average predicted earnings, and draws a fitted line for each gender/worker type across its corresponding values for the 10 clusters. If the $b(k)$ varied across clusters, signalling complementarities, this would result in differences in the slopes across worker types [10].

Overall, Figure 5 shows little evidence of complementarities both for males and females: for each gender, the slopes are similar across worker types, meaning that there is no differential benefit for different types of worker to be in a higher paying cluster relative to a lower paying one. Thus it is unlikely that match effects contribute meaningfully to the gender wage gap in our sample. To support this graphical evidence, we perform in Table 6 the decomposition of equation (6). The gender gap in $a(k)$, i.e. in average firm pay premiums without accounting for complementarities, is 0.044, very close to our additive model estimates, while what we call the gender gap in complementarities (term (9) of decomposition (6)) is basically zero and, if anything, negative.

Finally, Figure 6 reports the distribution of worker types across clusters for men and women. All types of workers are represented in each cluster but higher type workers are much more likely to be in higher paying clusters and lower type workers in lower paying ones. However, we see that the patterns are fairly similar across gender suggesting that, even though women are more likely to sort into lower paying clusters, the assortative matching between higher types workers and higher paying firms is not meaningfully different for men and women.

[10] Note, however, that the slopes on the figures are not directly interpretable as the $b(k)$ since, following BLM, the x-axis correspond to clusters, not to α_l^G.

5 Conclusion

We use a cluster-based approach proposed by Bonhomme et al. (2019) to further our understanding of the role of firms in explaining the gender wage gap. Using clusters allows us to take into account the universe of firms, to document changes in the role of firms over time and over the life cycle, and to account for potential complementarities between worker types and firms. Our findings suggest that lifting the dual connected set restriction imposed by the AKM methodology, and therefore better representing smaller firms, leads to higher estimates of the contribution of firms to the gender wage gap, driven by a higher within component. They also reveal that the gender gap in firm pay premiums has been roughly stable over the last twenty years but represents an ever increasing share of the unconditional gender wage gap. The decomposition into a between and within component has been fairly stable as well. The gender gap in firm pay premiums increases over the life-cycle due to increased sorting of women into lower paying firms as workers age. Finally we find limited evidence of complementarities for both men and women.

References

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High wage workers and high wage firms," *Econometrica*, 67, 251–333.

ANDREWS, M. J., L. GILL, T. SCHANK, AND R. UPWARD (2012): "High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias," *Economics Letters*, 117, 824–827.

BECKER, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846.

BLAU, F. D. AND L. M. KAHN (2017): "The Gender Wage Gap: Extent, Trends, and Explanations," *Journal of Economic Literature*, 55, 789–865.

BONHOMME, S., T. LAMADON, AND E. MANRESA (2019): "A distributional framework for matched employer employee data," *Econometrica*, 87, 699–739.

BRUNS, B. (2019): "Changes in Workplace Heterogeneity and How They Widen the Gender Wage Gap," *American Economic Journal: Applied Economics*, 11, 74–113.

CARD, D., A. R. CARDOSO, AND P. KLINE (2016): "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 131, 633–686.

CASARICO, A. AND S. LATTANZIO (2019): "What Firms Do: Gender Inequality in Linked Employer-Employee Data," WorkINPS Papers n. 24.

COSTA DIAS, M. AND K. HOLZHEU (2021): "Missing out on growth ? The Dynamic Gender Wage Gap in Firm Opportunities," .

COUDIN, E., S. MAILLARD, AND M. TÔ (2018): "Family, firms and the gender wage gap in France," IFS Working Papers W18/01, Institute for Fiscal Studies.

DI ADDARIO, S., P. KLINE, R. SAGGIO, AND M. SØLVSTEN (2021): "It Ain't Where You're From, It's Where You're At: Hiring Origins, Firm Heterogeneity, and Wages." Working Paper 28917, National Bureau of Economic Research.

EECKHOUT, J. AND P. KIRCHER (2011): "Identifying sorting-In theory," *Review of Economic Studies*, 78, 872–906.

Gallen, Y., R. V. Lesner, and R. Vejlin (2019): "The labor market gender gap in Denmark: Sorting out the past 30 years," *Labour Economics*, 56, 58–67.

Hagedorn, M., T. H. Law, and I. Manovskii (2017): "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 85, 29–65.

Morchio, I. and C. Moser (2020): "The Gender Pay Gap: Micro Sources and Macro Consequences," .

Olivetti, C. and B. Petrongolo (2016): "The Evolution of Gender Gaps in Industrialized Countries," *Annual Review of Economics*, 8, 405–434.

Shimer, R. and L. Smith (2000): "Assortative matching and search," *Econometrica*, 68, 343–369.

Sorkin, I. (2017): "The Role of Firms in Gender Earnings Inequality: Evidence from the United States," *American Economic Review*, 107, 384–87.

Tables and Figures

Table 1: Selected estimates of the contribution of firm pay policies to the gender wage gap

	(2) **Data source**	(3) **Wage measure**	(4) **Gender wage gap**	(5) **Gender gap in firm pay premiums**	(6) **Within (Bargaining)**	(7) **Between (Sorting)**
Card et al. (2016)	Portugal (2002-2009)	hourly	0.23	0.050	0.003	0.047
Bruns (2019)	West Germany (1994-2001)	daily	0.25	0.028	-0.014	0.042
	West Germany (2001-2008)	daily	0.25	0.064	0.001	0.063
Casarico & Lattanzio (2018)	Italy (1995-2015)	weekly	0.21	0.065	0.021	0.044
Coudin et al. (2018)	France (1995-2015)	hourly	0.17	0.014	-0.004	0.018
Morchio & Moser (2020)	Brazil (2007-2014)	hourly	0.14	0.084	0.020	0.064

Note: The table shows estimates of papers which report both the between and within component of the gender gap in firm pay premiums. Column 4 reports the gender gap in log wages of the estimation sample. Column 5 shows the difference between average firm effects for males and average firm effects for females while columns 6 and 7 show the split between a within and a between components (using female effects for the between term and the male distribution for the within one, see equation 3).

Table 2: Summary statistics

	Full sample		Dual Connected Set	
	Males	**Females**	**Males**	**Females**
Mean Age	40.41	40.17	40.38	39.91
Mean Tenure (years)	5.87	5.61	5.89	5.59
Mean Experience (years)	15.75	15.30	15.78	15.14
Mean Annual Earnings [a]	32,247	25,852	35,698	27,988
Mean Hourly Wage [b]	18.82	16.36	20.68	17.50
Share Part time	7.50%	27.00%	8.00%	27.10%
Mean Firm Size (# employees)	30.85		282.67	
Mean VA / worker	66.96		76.93	
Share of Female at the Firm (%)	39%		45%	
Number of worker-year obs	2,644,368	1,598,559	1,197,605	794,874
Number of workers	521,072	329,425	236,706	163,622
Number of firms	308,843		17,670	

[a] Gross, in Euros

[b] Gross, in Euros, in the main job

Note: The dual connected set refers to the intersection between the male and the female largest connected sets. The male/female largest connected sets are the relevant estimation sample when estimating equation (2) with firm dummies for the period 2009-2015 separately for males and females. The full sample is the estimation sample when using cluster dummies in equation (2) for the period 2009-2015.

Table 3: Comparing average gender-specific cluster and firm effects, holding fixed normalization and sample

Specification	$E(\Psi_k^M \mid Male)$	$E(\Psi_k^M \mid Female)$	$E(\Psi_k^F \mid Male)$	$E(\Psi_k^F \mid Female)$
Firm dummies	0.107	0.088	0.098	0.078
10 Cluster dummies	0.119	0.093	0.112	0.088

Note: Average firm or cluster effects for males or females come from estimating equation (2), separately for each gender, with either firm or clusters dummies. Normalization is done with respect to low value added firms. The sample is the dual connected set for 2009-2015.

Table 4: Decomposition of the gender gap in pay premiums: firms v. clusters within the dual connected set

Specification	Gender gap in firm pay premium	Between component	Within component
Firm Dummies	0.029	0.020	0.009
		69%	31%
10 clusters	0.031	0.024	0.007
		77%	23%
50 clusters	0.032	0.024	0.008
		75%	25%
200 clusters	0.032	0.024	0.008
		75%	25%

Note: We use the cluster or firm effects estimated from equation (2), restricting attention to firms in the dual connected set of 2009-2015 and normalizing with respect to low value added firms. We show results from decomposition (3).

Table 5: Decomposition of the gender gap in pay premiums:
dual connected set v. full sample

Sample	Gender wage gap	Gender gap in firm pay premium	Between component	Within component
Dual Connected Set	0.170	0.031	0.024 77%	0.007 23%
Full Sample	0.135	0.042	0.021 50%	0.021 50%

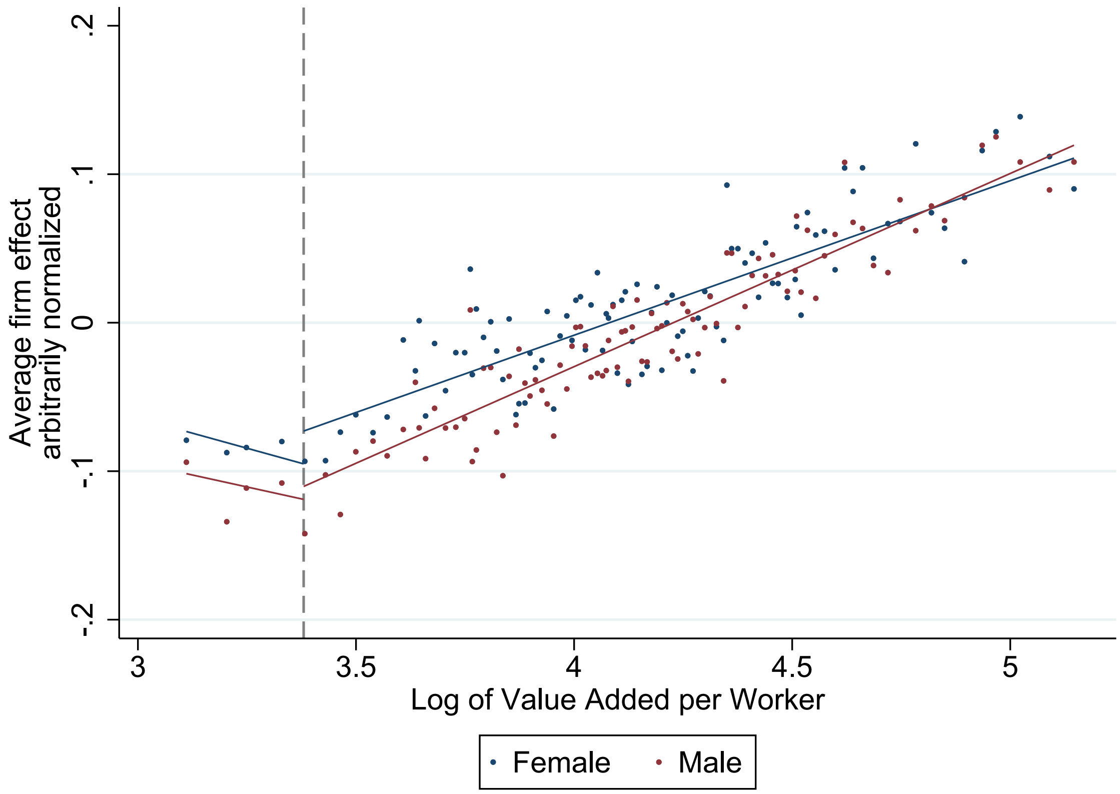
Note: We show results from the decomposition (3), using the cluster effects estimated from equation (2), with 10 clusters and normalizing with respect to low value added firms. The first row restricts attention to the dual connected set of 2009-2015 while the last row uses the full sample for that same period.

Table 6: Decomposition of the gender wage gap - Complementarities

Sample	Specification	Gender wage gap	Gender gap in firm pay premiums	Gender gap in complementarities
Full sample	Model with interaction	0.135	0.044	-0.004

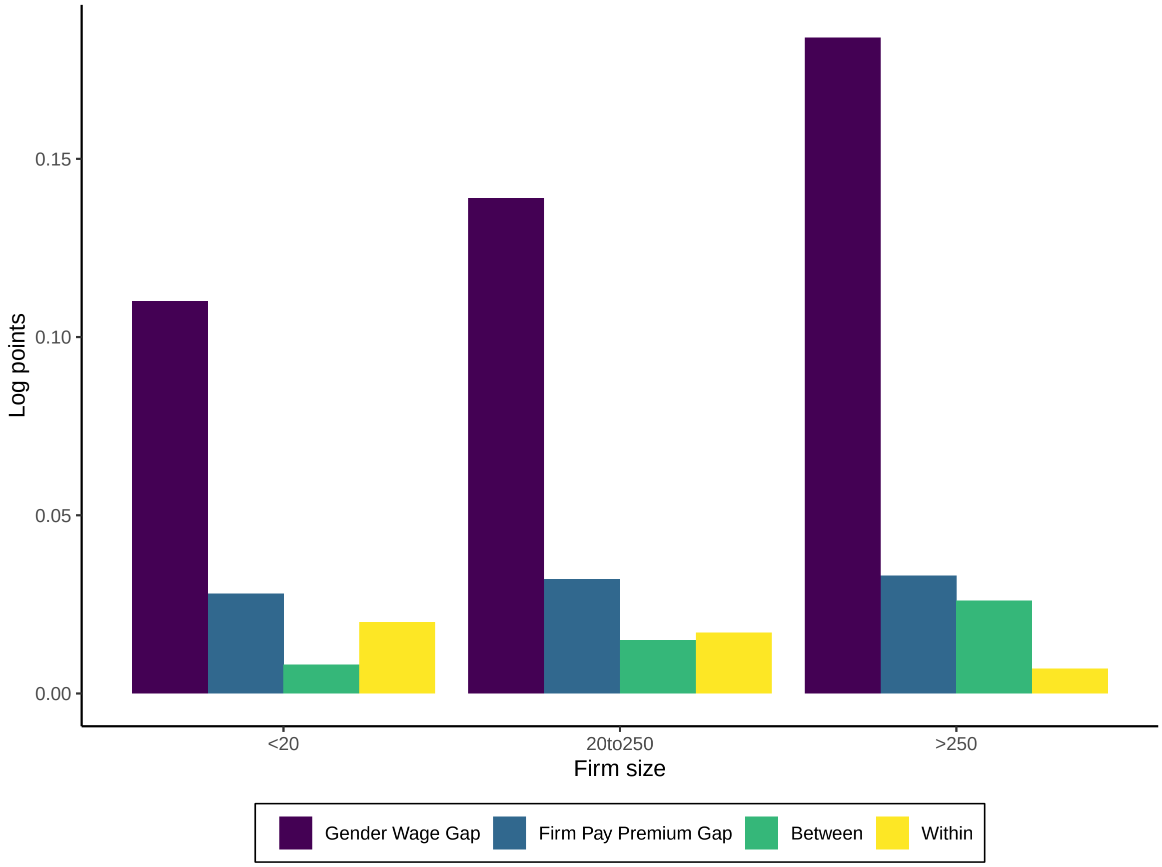
Note: We use the cluster effects estimated from equation (5), normalized with respect to low value added firms, and show results from decomposition (6). The gender gap in firm pay premiums (column 4) corresponds to term (7) and the gender gap in complementarities (column 5) corresponds to term (9).

Figure 1: Average firm fixed effects by bins of value added per worker



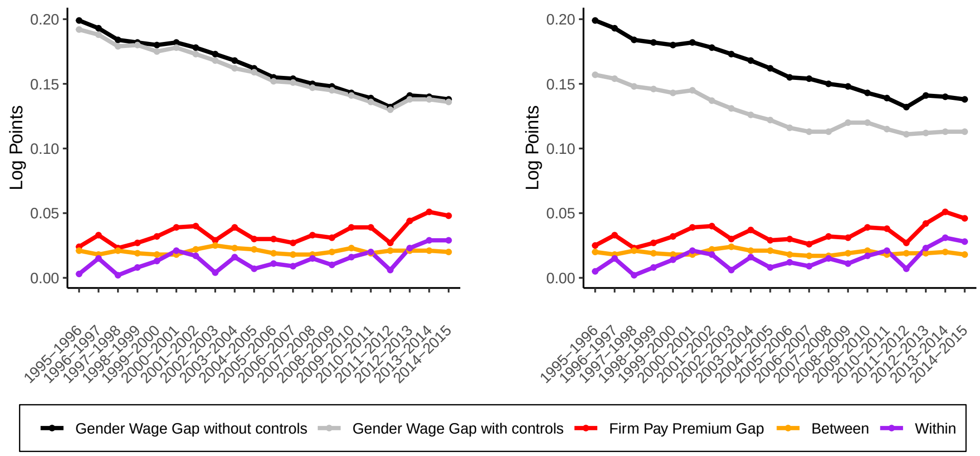
Note: The figure shows the average estimated male firm effects (in red) and the average estimated female firm effects (in blue) for each bin of value added per worker. These bins are computed employment weighted. Firm effects are estimated using equation (2).

Figure 2: Heterogeneity of the contribution of firms to the gender wage gap by firm size



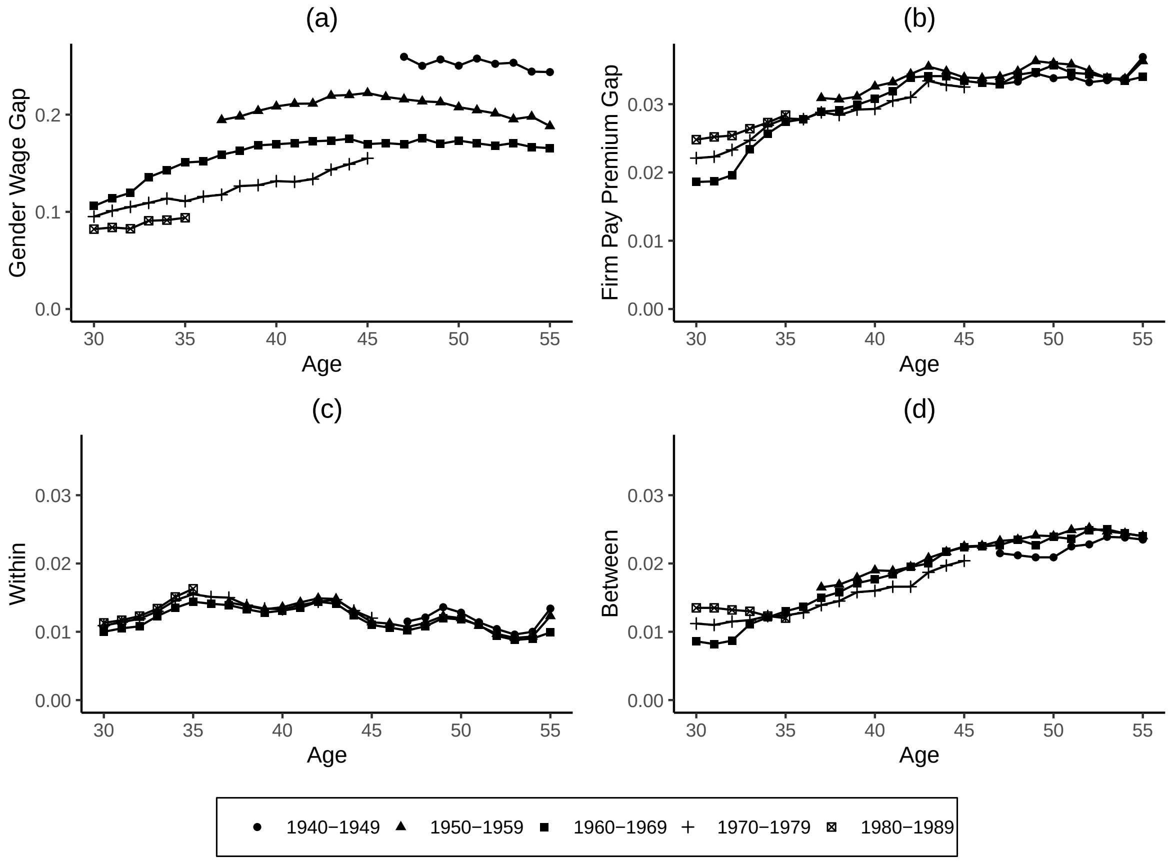
Note: This Figure uses our main estimates (using 10 clusters, the full sample and normalizing with respect to value added, reported in Table 5) but splits the sample into 3 based on firm size, measured by the number of employees. It then reports the average unconditional gender wage gap, the average gender gap in pay premiums and its decomposition into a between and within components as described in equation (3) for each of the 3 sub-samples.

Figure 3: Evolution of the contribution of firms to the gender wage gap over time



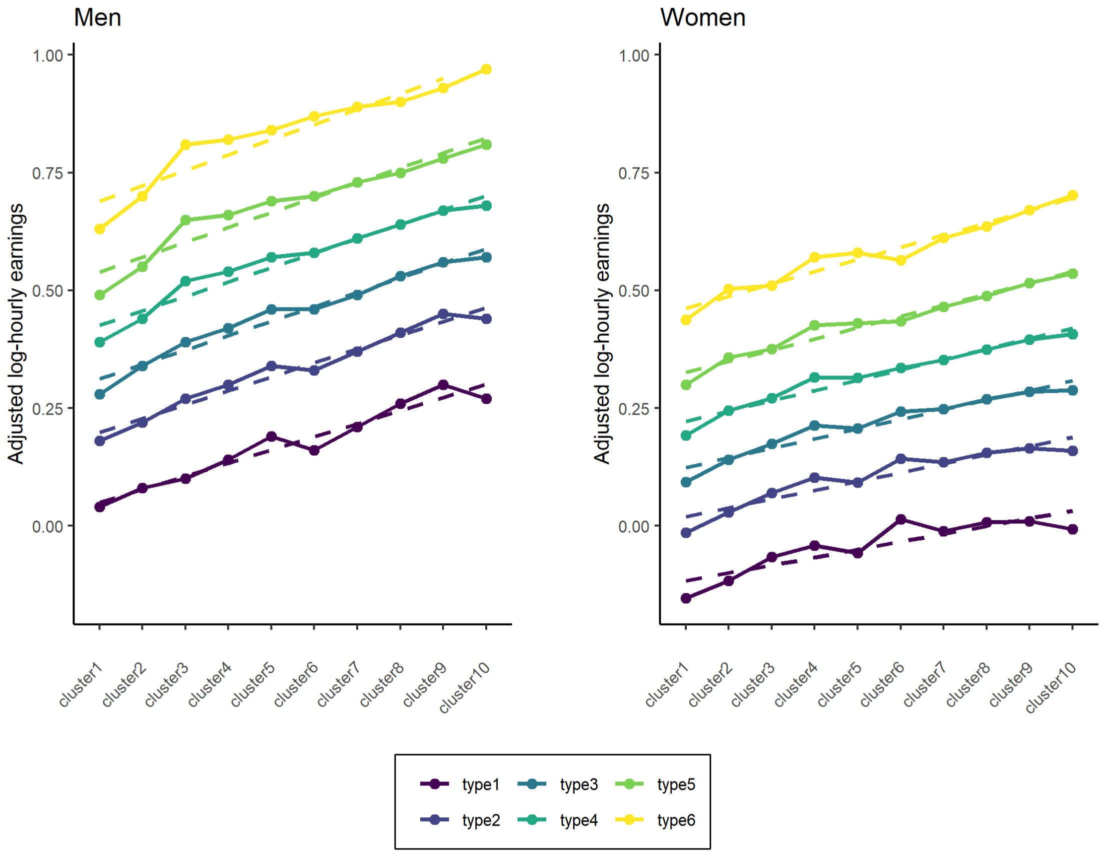
Note: The black line gives the unconditional gender gap in hourly wage. The grey line gives the gender gap in hourly wage controlling for the covariates used when estimating the firm/cluster effects in equation (2), i.e. a polynomial in age, experience and tenure in the left panel as well as 2-digit occupations in the right panel. The red, orange and purple lines give the gender gap in firm pay premiums and and its decomposition into a between and within components following equation (3), using 10 clusters and normalizing with respect to low value added firms.

Figure 4: Evolution of the contribution of firms to the gender wage gap over the life-cycle



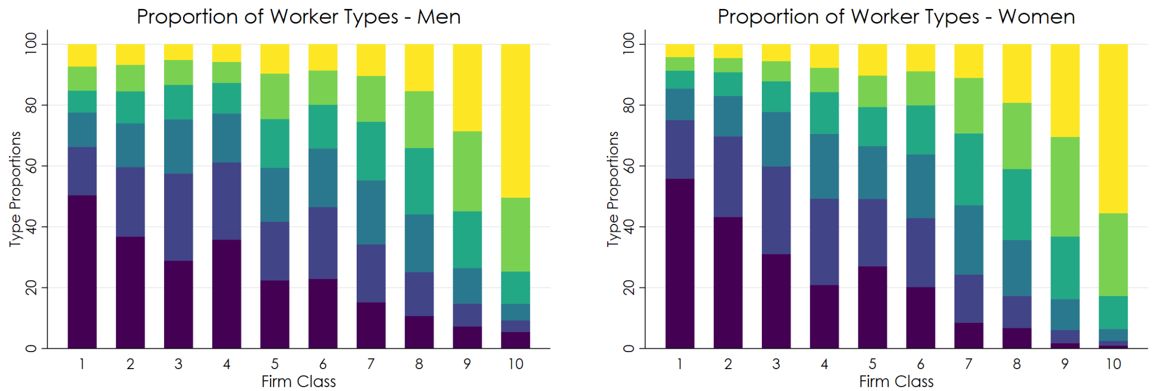
Note: This Figure reports, by age and cohort, estimates of the unconditional gender wage gap (panel a), of the gender gap in firm pay premiums (panel b) and of its decomposition into a within (panel c) and between (panel d) components. We estimate equation (2) on overlapping 5 years age brackets. For each age bracket, we normalize the gender specific cluster effects relative to low value added firms. We then compute averages of these estimates by gender, age and cohort and report results of decomposition (3). For each specific age, the value shown on the figure corresponds to the average value for the 5 brackets that include it.

Figure 5: Estimates of complementarities, by gender



Note: This shows graphically estimates of the model of equation (5). We order clusters from the lowest in terms of average wage to the highest. Colors refer to different types of workers, from low type workers (dark blue) to high types (yellow). For each worker type and cluster, the graph presents estimates of mean log hourly wages. Differences in slopes across worker types would be evidence in favor of complementarities.

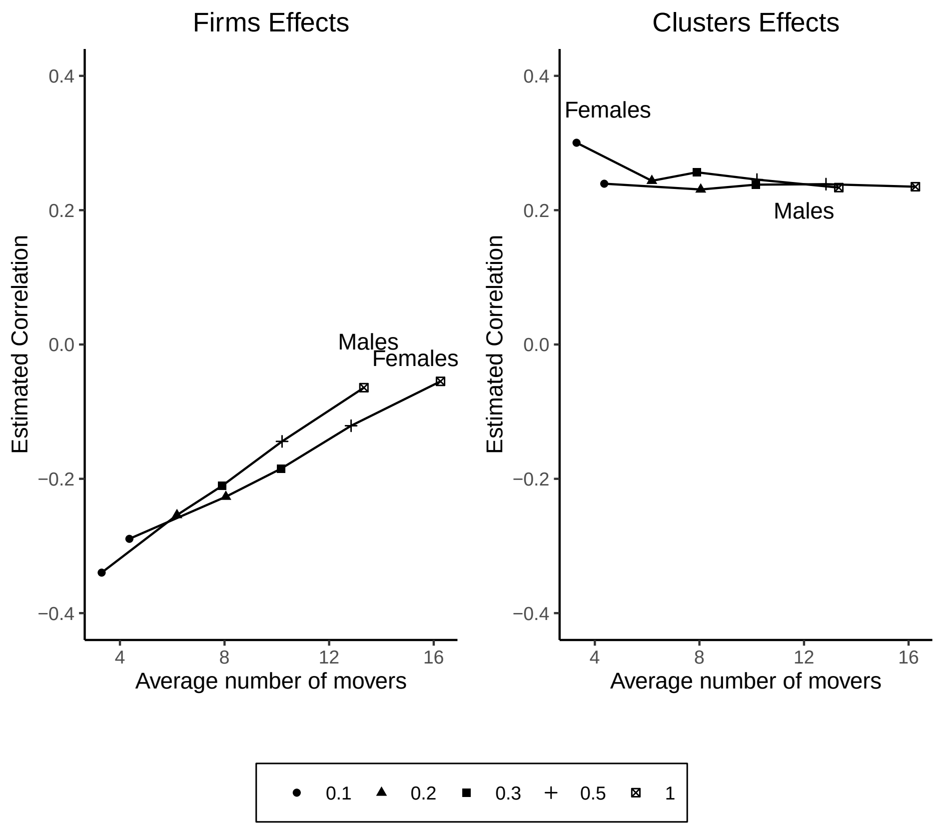
Figure 6: Distribution of workers of different types across the different clusters, by gender



Note: Colors refer to different types of workers, from low type workers (dark blue) to high types (yellow). We order clusters from the lowest in terms of average wage to the highest.

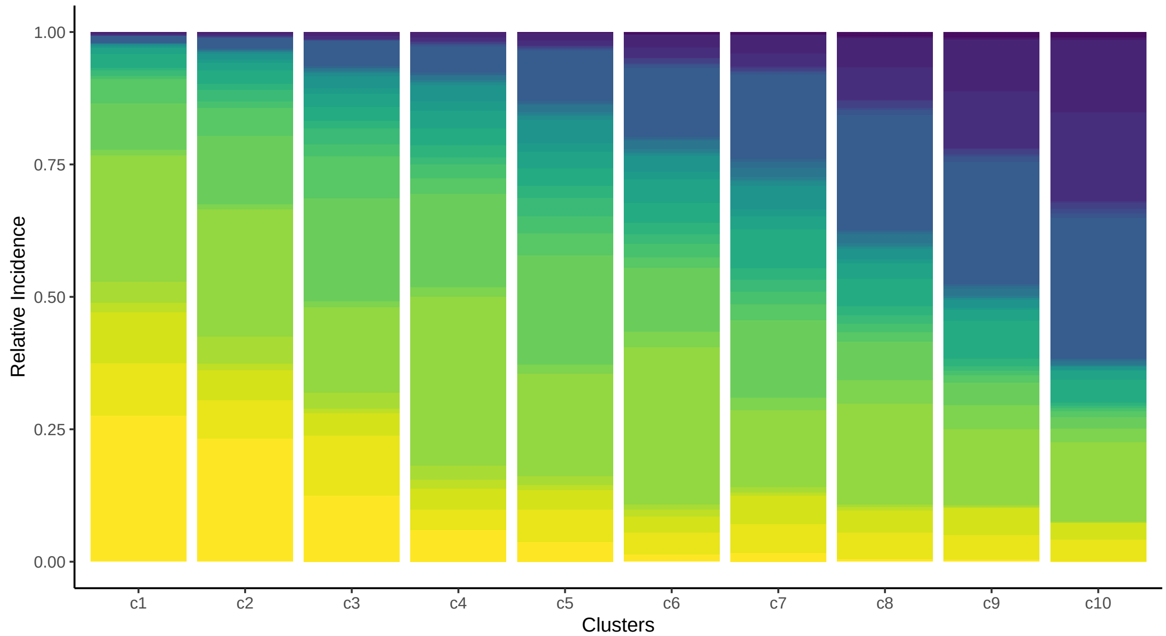
Appendix Figures

Figure A1: Correlation of individual and firm/cluster fixed effects, by gender, when increasing the number of workers per firm in a fixed sample of firms



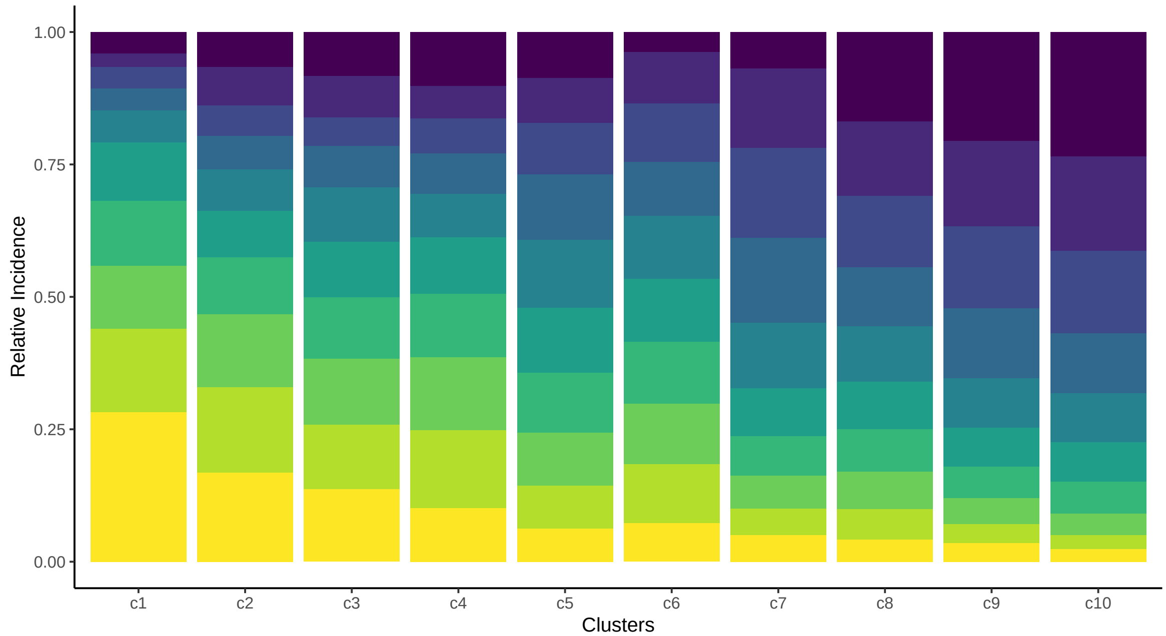
Note: This figure shows the evolution of the correlation between individual fixed effects and firm/cluster effects for the female and male sample increasing the number of workers (and hence movers) per firm. As in Andrews et al. (2012), we take a random sample of workers, record their firms, and calculate the correlation between the individual and firm or cluster fixed effects. We then increase the random sample of workers while holding the sample of firms in which they work fixed and recalculate the correlation between individual and firm or cluster fixed effects. This correlation, for samples of workers ranging from 10 % to 100 %, is negative and increasing with sample size in the left panel (using AKM methodology) while positive and relatively stable in the right panel (using the cluster approach). The sample is the dual connected set for the period 2009-2015.

Figure A2: Sector incidence by clusters



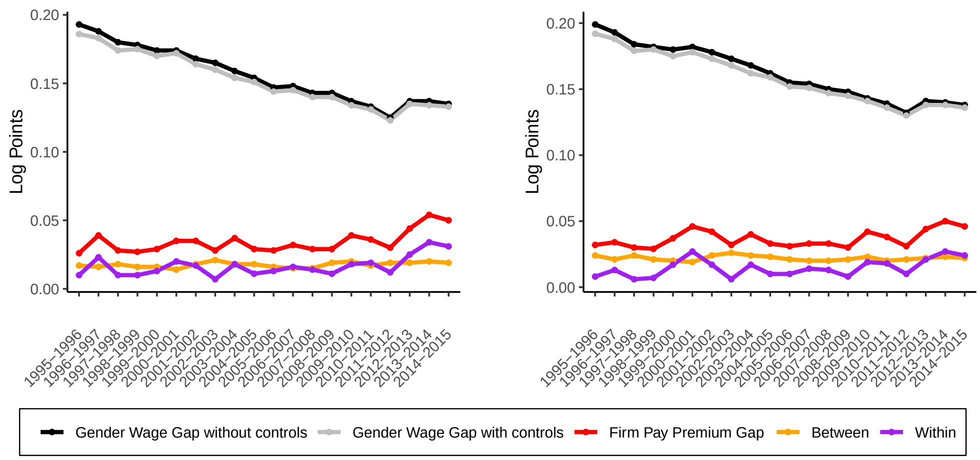
Note: This figure shows the relative incidence of sectors across clusters. We order clusters from the lowest in terms of average wage (c1) to the highest (c10). Sectors are classified according to the 2-digit nomenclature A38 from yellow-ish to blues-ish.

Figure A3: Firm size incidence by clusters



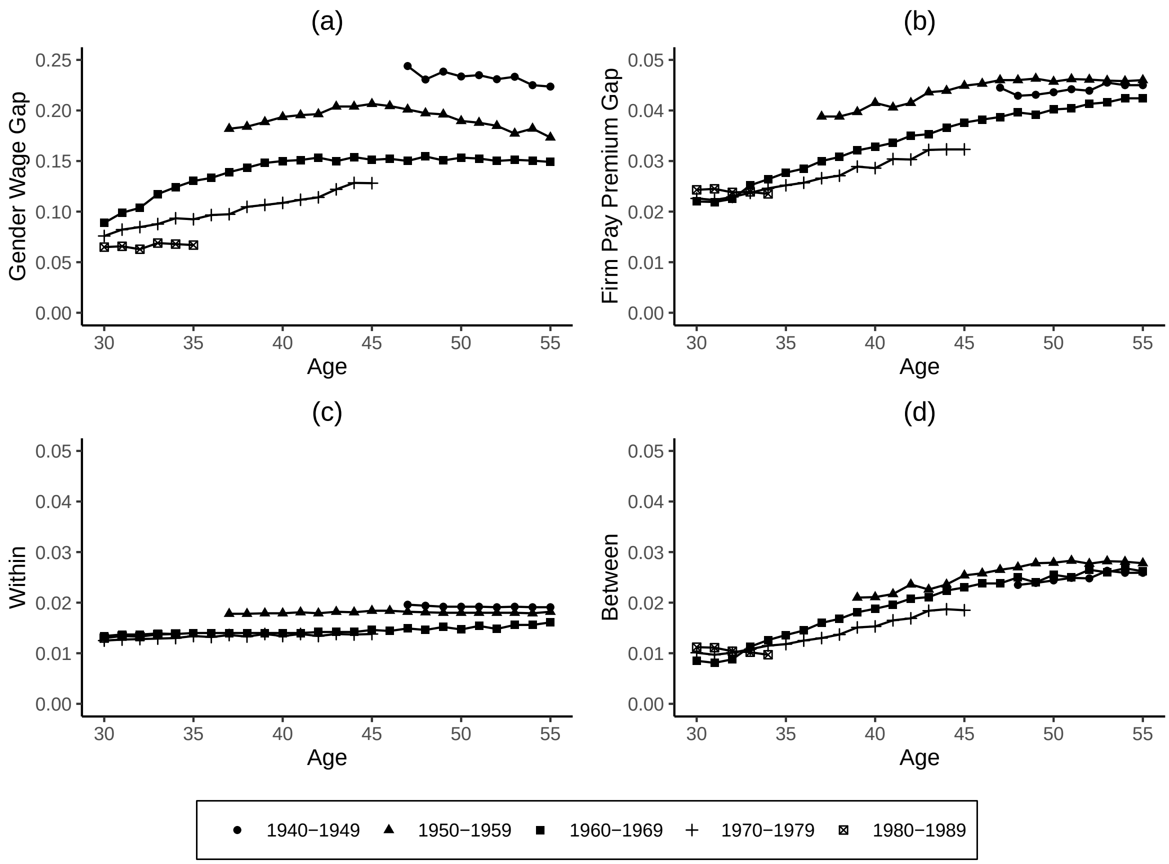
Note: This figure shows the relative incidence of firm size deciles across clusters. We order clusters from the lowest in terms of average wage (c1) to the highest (c10). As for firm size deciles: yellow-ish refers to smaller firms and blues-ish to bigger firms.

Figure A4: Evolution of the contribution of firms to the gender wage gap over time



Note: The black line gives the unconditional gender gap in hourly wage. The grey line gives the gender gap in hourly wage controlling for the covariates used when estimating the firm effects in equation (2), i.e. age, tenure and experience in our baseline. The red, orange and purple lines give the gender gap in firm pay premiums and and its decomposition into a between and within components following equation (3). The left panel shows robustness to normalizing with respect to the hotel and restaurant sector. The right panel shows robustness to using 50 clusters (and still normalizing with respect to low value-added firms as in our benchmark results of Figure 3).

Figure A5: Evolution of the contribution of firms to the gender wage gap over the life-cycle



Note: This Figure reports, by age and cohort, estimates of the unconditional gender wage gap (panel a), of the gender gap in firm pay premiums (panel b) and of its decomposition into a within (panel c) and between (panel d) components. We estimate equation (2) by cohort. For each cohort, we normalize the gender specific cluster effects relative to low value added firms. We then compute averages of these estimates by gender, age and cohort and report results of decomposition (3).

Appendix Tables

Table A1: Variance Decomposition - Models without complementarities

Firm effect	Sample		$Var(w)$	$\frac{Var(\alpha)}{Var(w)}$	$\frac{Var(\Psi)}{Var(w)}$	$\frac{2Cov(\alpha,\Psi)}{Var(w)}$	$\frac{Var(r)}{Var(w)}$
AKM	DC	Male	0.1603	80.23%	16.98%	-4.99%	7.78%
	DC	Female	0.1449	77.93%	20.58%	-7.94%	9.43%
10 Clusters	DC	Male	0.1603	71.43%	4.12%	15.94%	8.56%
	DC	Female	0.1449	65.97%	4.75%	19.15%	10.52%
10 Clusters	Full Sample	Male	0.1572	67.55%	5.16%	16.29%	11.00%
	Full Sample	Female	0.1382	63.98%	5.38%	18.26%	12.38%

Note: This table shows variance decomposition of log hourly wages using different samples for the period 2009-2015 and the additive model of equation (2). DC refers to the dual connected set. w denotes the log hourly wage, α the individual fixed effects, Ψ the firm/cluster effects and r the residuals.

Table A2: Variance Decomposition - Model with complementarities

Firm effect	Sample		$\frac{Var(\alpha)}{Var(Y)}$	$\frac{Var(\Psi)}{Var(Y)}$	$\frac{2Cov(\alpha,\Psi)}{Var(Y)}$
10 clusters	Full sample	Male	78.96%	3.94%	17.10%
10 clusters	Full sample	Female	77.60%	3.33%	19.06%

Note: This table shows variance decomposition of log hourly wages using the static model with interaction terms of equation (5) for the period 2009-2015. Y denotes the residualized log hourly wage, α the individual random effects and Ψ the cluster effects. The output comes from the BLM code made available online.